Two-Sided A Posteriori Error Estimates for Mixed Formulations of Elliptic Problems

Sergey Repin* Stefan Sauter† Anton Smolianski‡

Abstract

The present work is devoted to the a posteriori error estimation for mixed approximations of linear self-adjoint elliptic problems. New guaranteed upper and lower bounds for the error measured in the natural product norm are derived, and the individual sharp upper bounds are obtained for approximation errors in each of the physical variables. All estimates are reliable and valid for any approximate solution from the class of admissible functions. The estimates contain only global constants depending solely on the domain geometry and the given operators. Moreover, it is shown that, after an appropriate scaling of the coordinates and the equation, the ratio of the upper and lower bounds for the error in the product norm never exceeds 3. The possible methods of finding the approximate mixed solution in the class of admissible functions are discussed. The estimates are computationally very cheap and can also be used for the indication of the local error distribution. As the applications, the diffusion problem as well as the problem of linear elasticity are considered.

Keywords: a posteriori estimate, two-sided bounds, mixed approximation, elliptic problem

1 Introduction

Most of the existing elliptic problems of continuum mechanics are originally derived in the mixed form, i.e. they contain *two* physical variables that are often equally important in the applications. For example, the stationary heat conduction (resp., diffusion) equation

$$-\Delta u + f = 0$$

comes from the energy (resp., mass) balance equation

$$-\operatorname{div} \mathbf{p} + f = 0 \tag{1}$$

and the empirical Fourier (resp., Fick) law for the heat (resp., mass) flux

$$\mathbf{p} = \nabla u\,. \tag{2}$$

Here we have set, for simplicity, the conduction (diffusion) coefficient equal to 1 and changed the sign in the flux relation (2). Both the temperature (molecular concentration)

*V.A. Steklov Institute of Mathematics, Fontanka 27, 191 011 St. Petersburg, Russia; E-mail: repin@pdmi.ras.ru

†Institute of Mathematics, Zurich University, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland; E-mail: stas@math.unizh.ch

‡Institute of Mathematics, Zurich University, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland; E-mail: antsmol@amath.unizh.ch

u and the flux $\mathbf{p}$ may be needed for understanding the real physical process, and this requirement becomes of utmost importance in the problems of the flows in porous media and in the elasticity problems, where the *complete* solution of the problem is the *pair* of pressure/velocity, respectively, displacement/stress variables.

These considerations served as a motivation for the extensive research in the field of the so-called mixed methods, that is the methods allowing to obtain the approximations to the both physical variables of the problem. As general references on the subject the books [9], [29] and [17] can be recommended. Although the computing methods are very important, it is also required for reliable modelling to have an explicit error control for the obtained approximations. This issue, namely, the *a posteriori* error estimation for the mixed formulations of elliptic problems constitutes the primary goal of the present work.

There have been quite a few papers on the a posteriori error estimation for the mixed finite element methods (FEM). The residual-based estimates were developed in [2], [6], [12], [1], [16] for the diffusion-type equation and extended in [14] and [20] to the equations of linear elasticity. The superconvergence-based (averaging-type) error estimators were proposed in [7] and [13] to control the L_2-error of the flux variable. Further, the estimators based on the solution of local problems were presented in [2], [16] and [20], and the hierarchical estimator can be found in [31]. Finally, a comparison of these four types of error estimators for mixed finite element discretizations by Raviart-Thomas elements was presented in [31].

In this paper, we derive a posteriori error estimates of another type, the so-called *functional-type estimates* (see also [24], [25], [26], [27]). For the example of problem (1)–(2) equipped with the zero Dirichlet boundary condition for u and under the assumption that $f \in L_2(\Omega)$ (where Ω is the physical domain), the main estimates look as follows:

$$\begin{aligned}
\|(u-v, \mathbf{p}-\mathbf{y})\|_{1\times\mathrm{div}} &\leq \|\nabla v - \mathbf{y}\| + (1+2C_\Omega^2)^{1/2}\|\mathrm{div}\,\mathbf{y} - f\|\,, \\
\|(u-v, \mathbf{p}-\mathbf{y})\|_{1\times\mathrm{div}} &\geq \frac{1}{\sqrt{3}}\left(\|\nabla v - \mathbf{y}\| + \|\mathrm{div}\,\mathbf{y} - f\|\right)\,.
\end{aligned}$$

Here $\|\cdot\|$ is the norm in $L_2(\Omega)$, the *full* norm $\|(\cdot,\cdot)\|_{1\times\mathrm{div}}$ is defined as

$$\|(u-v, \mathbf{p}-\mathbf{y})\|_{1\times\mathrm{div}} := (\|\nabla(u-v)\|^2 + \|\mathbf{p}-\mathbf{y}\|^2 + \|\mathrm{div}\,(\mathbf{p}-\mathbf{y})\|^2)^{1/2}\,,$$

and the pair $(v, \mathbf{y})$ from the product space $H_0^1(\Omega) \times H(\Omega; \mathrm{div})$ is *any* approximate solution to the mixed problem. The constant C_Ω is the global constant from the Friedrichs inequality and depends only on the domain geometry.

We see that, while these estimates provide *guaranteed* upper and lower bounds for the error of the mixed solution in the *full* norm, the estimates are also very flexible in the sense that they can be applied to a variety of different approximations, not being restricted to a particular discretization method. This fact makes the functional-type estimates especially attractive for the control of the modelling errors, like those arising in dimension reduction methods of continuum mechanics (see [28]). The sharpness of the estimates and the ability to indicate the local error distribution required for the mesh adaptation will also be shown. Last but not least, we remark that, once the approximate solution has been found in the product space, the estimates cost very little: their computation amounts to the calculation of the corresponding norms.

It is worth noting that, if the given data $f \in L_2(\Omega)$, the exact mixed solution $(u, \mathbf{p})$ belongs to the product space $H^1(\Omega) \times H(\Omega; \mathrm{div})$, which means that this space is the natural space for the approximation of the mixed solution. While the standard primal and dual mixed FEM approximate the mixed pair in $H^1(\Omega) \times L_2(\Omega; \mathbb{R}^n)$, resp., $L_2(\Omega) \times H(\Omega; \mathrm{div})$ (hence, not using the full regularity of the exact solution), there are alternative methods that allow to contsruct the approximate solution directly in $H^1(\Omega) \times H(\Omega; \mathrm{div})$. Some of these methods seem to be very promising and competitive, also in the case when one wants to find an approximation of the flux (stress) variable only. Although the comparative analysis of these methods is a subject of the next paper, we briefly review here four of them, since it is important for the application of our error estimates.

The rest of the paper is organized as follows. In Section 2, we introduce the notation for the mixed formulation of a general linear self-adjoint elliptic problem. In Section 3, the two-sided sharp a posteriori error estimate is derived for an arbitrary approximate solution from the natural class of admissible functions. Next, the individual a posteriori estimates for each of the two variables are derived and shown to be sharp as well. Section 4 is devoted to the applications of the developed theory. First, we consider the diffusion problem and obtain the explicit error bounds for its approximate mixed solution, then, we discuss the possible methods of constructing the solution in the natural product space. Finally, the a posteriori error estimates are derived for both displacement and stress approximations in the problem of linear elasticity.

2 Preliminaries

Let V be a reflexive Banach space with the norm $\|\cdot\|_V$, Y a Hilbert space equipped with the inner product $(\cdot,\cdot)_Y$ and the norm $\|\cdot\|_Y$, V_0 a linear subspace of V. By $\mathcal{B}$ we denote a linear bounded operator acting from V into Y, by $\mathcal{B}^* : Y \to V_0^*$ the dual operator to $\mathcal{B}\big|_{V_0}$ (the restriction of $\mathcal{B}$ to V_0) in the sense that, for any $y \in Y$,

$$(y, \mathcal{B}w)_Y = \langle \mathcal{B}^* y, w \rangle \quad \forall w \in V_0 \,.$$

Here $\langle w^*, w \rangle$ denotes the value of the functional $w^* \in V_0^*$ on the element $w \in V_0$.

Next, let us introduce a self-adjoint operator $\mathcal{A} \in \mathcal{L}(Y,Y)$ such that

$$\lambda_A \|y\|_Y^2 \;\le\; (\mathcal{A}y, y)_Y \;\le\; \Lambda_A \|y\|_Y^2 \quad \forall y \in Y \,, \tag{3}$$

where λ_A and Λ_A are positive constants independent of y. Such an operator defines the equivalent norm on Y

$$|\!|\!|\, y \,|\!|\!| := (\mathcal{A}y, y)_Y^{1/2} \,.$$

The inverse operator $\mathcal{A}^{-1}$ satisfies an inequality of type (3) with constants Λ_A^{-1} and λ_A^{-1} and defines another equivalent norm on Y

$$|\!|\!|\, y \,|\!|\!|_* := (\mathcal{A}^{-1}y, y)_Y^{1/2} \,.$$

Assume also that the operator $\mathcal{B}$ satisfies the coercivity inequality on V_0

$$\|w\|_V \;\le\; C_{\mathcal{B}} \|\mathcal{B}w\|_Y \quad \forall w \in V_0 \,, \tag{4}$$

where $C_{\mathcal{B}}$ is some positive constant independent of w. Using (3) and (4) one can define an equivalent norm $|\!|\!|\, \mathcal{B}\cdot |\!|\!|$ on V_0 as well as the following norm on the dual space V_0^*:

$$[\![\, w^* \,]\!] := \sup_{w \in V_0 \setminus \{0\}} \frac{\langle w^*, w \rangle}{|\!|\!|\, \mathcal{B}w \,|\!|\!|} \,.$$

Let now u_0 be some given function from V and

$$V_0 + u_0 := \{ v \in V \;\mid\; v = w + u_0 \,,\; w \in V_0 \} \,.$$

Let, in addition, l be some given functional from V_0^*. Then, the problem
$\underline{(\mathcal{P})}$: Find $u \in V_0 + u_0$ such that

$$(\mathcal{A}\mathcal{B}u, \mathcal{B}w)_Y + \langle l, w \rangle = 0 \quad \forall w \in V_0$$

has the unique solution (this follows from (3) and (4)).

The problem can be rewritten in the operator form as follows:

$$\mathcal{B}^* \mathcal{A}\mathcal{B}u + l = 0 \quad \text{in } V_0^* .$$

The mixed formulation of the problem can be immediately obtained by the introduction of the new unknown function

$$p = \mathcal{A}\mathcal{B}u \,,$$

which leads to the problem
$\underline{(\mathcal{M})}$: Find $(u, p) \in (V_0 + u_0) \times Y$ such that

$$p = \mathcal{A}\mathcal{B}u \quad \text{in } Y \,,$$
$$\mathcal{B}^* p + l = 0 \quad \text{in } V_0^* .$$

It is clear that the problem $(\mathcal{M})$ has the well known saddle-point structure; its unique solvability is a direct consequence of conditions (3) and (4) (see, e.g., [9]).

In the sequel, we will adopt the terminology used in the duality theory of convex analysis (see, e.g., [15]) and call the solution of the problem $(\mathcal{P})$ the *primal variable* (*primal solution*) and the new unknown p the *dual variable* (*dual solution*). Accordingly, the letters u, v, w will be reserved for the functions related to the primal variable, i.e. belonging to the space V, whilst the letters p, q, y for those related to the dual variable, i.e. being in the space Y.

In view of the second equation of the problem $(\mathcal{M})$, the dual variable p belongs to the set

$$Q_l := \{q \in Y \mid \mathcal{B}^* q = -l \text{ in } V_0^*\} .$$

Thus, for the full control of the dual variable one needs an extended norm on Y, that we define as

$$\|y\|_{\mathcal{B}^*} := \left(||| \, y \, |||_*^2 + [\![\mathcal{B}^* y]\!]^2 \right)^{1/2} \quad \forall y \in Y . \tag{5}$$

Although this is an equivalent norm on Y (since $[\![\mathcal{B}^* y]\!] = \sup\limits_{w \in V_0} \frac{\langle \mathcal{B}^* y, w \rangle}{|||\mathcal{B}w|||} = \sup\limits_{w \in V_0} \frac{(y, \mathcal{B}w)_Y}{|||\mathcal{B}w|||} \leq \sqrt{\frac{\Lambda_A}{\lambda_A}} \, ||| \, y \, |||_*$), we need it to explicitly control the error in the "equilibrium equation", i.e. in the second equation of $(\mathcal{M})$.

Finally, we define the *full norm* on the product space $V_0 \times Y$:

$$\|(w, y)\|_{V_0 \times Y} := \left(||| \, \mathcal{B}w \, |||^2 + \|y\|_{\mathcal{B}^*}^2 \right)^{1/2} \quad \forall (w, y) \in V_0 \times Y . \tag{6}$$

3 General estimates

3.1 Estimate in the full norm

Let $(v, q) \in (V_0 + u_0) \times Q_l$ be an arbitrary approximation to the exact solution (u, p) of the problem $(\mathcal{M})$. Then, with the help of the relation $p = \mathcal{A}\mathcal{B}u$ and the fact that $\mathcal{A}$ is a self-adjoint linear operator, it is easy to show (see also [25]) that

$$\begin{aligned} ||| \, \mathcal{B}(u - v) \, |||^2 + ||| \, p - q \, |||_*^2 &= (\mathcal{A}\mathcal{B}(u - v), \mathcal{B}(u - v))_Y + (\mathcal{A}^{-1}(p - q), p - q)_Y \\ &= (\mathcal{A}\mathcal{B}v - q, \mathcal{B}v - \mathcal{A}^{-1} q)_Y + 2(p - q, \mathcal{B}(u - v))_Y . \end{aligned}$$

Since both p and q belong to the set Q_l, $\mathcal{B}^*(p-q) = -l+l = 0$ in V_0^* and $(p-q, \mathcal{B}(u-v))_Y = \langle \mathcal{B}^*(p-q), u-v \rangle = 0$; this implies

$$\| \mathcal{B}(u-v) \|^2 + \| p-q \|_*^2 = (\mathcal{A}\mathcal{B}v - q, \mathcal{B}v - \mathcal{A}^{-1}q)_Y = \| \mathcal{A}\mathcal{B}v - q \|_*^2 \ . \qquad (7)$$

This equality can be referred to as the generalized Prager-Synge hypercircle identity (see [22]).

Relation (7) may already be viewed as an a posteriori error estimate, since the right-hand side does not depend on the exact solution (u, p). However, the estimate holds only for $q \in Q_l$, which seriously restricts the field of its practical application. In fact, the constraint of the set Q_l is virtually impossible to satisfy exactly (this would be nearly equivalent to finding the exact dual solution p), and that is why it is desirable to obtain an estimate allowing the approximate dual solution to be in some large *unconstrained* space.

If we waive the constraint $\mathcal{B}^*q = -l$ in V_0^* for the approximate dual variable, the latter remains to be considered in the whole space Y. Let $y \in Y$ be an arbitrary approximation to p, and $v \in V_0 + u_0$ as before. Then, using (7) one can derive

$$\begin{aligned} \| \mathcal{B}(u-v) \|^2 + \| p-y \|_*^2 &= \| \mathcal{B}(u-v) \|^2 + \| p-q \|_*^2 + 2(\mathcal{A}^{-1}(p-q), q-y)_Y \\ &+ \| q-y \|_*^2 = \| \mathcal{A}\mathcal{B}v - q \|_*^2 + 2(\mathcal{A}^{-1}(p-q), q-y)_Y + \| q-y \|_*^2 \quad \forall q \in Q_l \, . \qquad (8) \end{aligned}$$

Now we would like to estimate the right-hand side of (8) from above, so as to eliminate $q \in Q_l$. It is clear that, in order to obtain an explicitly computable and efficient upper bound, one has to carefully choose some special q in Q_l.

First, define the auxiliary function $w_y \in V_0$ such that

$$\mathcal{B}^*\mathcal{A}\mathcal{B}w_y = l + \mathcal{B}^*y \quad \text{in } V_0^* \, .$$

Due to assumptions (3) and (4) this problem has a unique solution.

Set now $q := y - \mathcal{A}\mathcal{B}w_y$. It is evident that such a function q belongs to Y and $\mathcal{B}^*q = \mathcal{B}^*y - \mathcal{B}^*\mathcal{A}\mathcal{B}w_y = \mathcal{B}^*y - l - \mathcal{B}^*y = -l$ in V_0^*, that is $q \in Q_l$. It may be noticed that, with this specific choice of q, the sum $q + \mathcal{A}\mathcal{B}w_y$ obviously becomes a non-orthogonal variant of the Helmholtz decomposition for the function $y \in Y$.

Now we can plug the constructed q into the right-hand side of (8). For the 1st term we have

$$\| \mathcal{A}\mathcal{B}v - q \|_* \leq \| \mathcal{A}\mathcal{B}v - y \|_* + \| \mathcal{A}\mathcal{B}w_y \|_* \ . \qquad (9)$$

Here $\| \mathcal{A}\mathcal{B}w_y \|_* = \left(\mathcal{A}^{-1}\mathcal{A}\mathcal{B}w_y, \mathcal{A}\mathcal{B}w_y\right)_Y^{1/2} = \| \mathcal{B}w_y \|$. The latter norm can be estimated by

$$\| \mathcal{B}w_y \|^2 = \langle \mathcal{B}^*\mathcal{A}\mathcal{B}w_y, w_y \rangle \ \leq \ [\![\mathcal{B}^*\mathcal{A}\mathcal{B}w_y]\!] \ \| \mathcal{B}w_y \| \, ,$$

which implies $\| \mathcal{B}w_y \| \leq [\![\mathcal{B}^*\mathcal{A}\mathcal{B}w_y]\!]$. We notice now that, by the definition of w_y, $\mathcal{B}^*\mathcal{A}\mathcal{B}w_y = l + \mathcal{B}^*y$ in V_0^*, that ultimately leads to the estimate

$$\| \mathcal{A}\mathcal{B}w_y \|_* \leq [\![l + \mathcal{B}^*y]\!] \, . \qquad (10)$$

Inserting this into (9) one obtains

$$\| \mathcal{A}\mathcal{B}v - q \|_* \leq \| \mathcal{A}\mathcal{B}v - y \|_* + [\![l + \mathcal{B}^*y]\!] \, . \qquad (11)$$

The 2nd term on the right-hand side of (8) can be rewritten as

$$\begin{aligned} 2(\mathcal{A}^{-1}(p-q), q-y)_Y = 2(\mathcal{A}^{-1}(p-q), -\mathcal{A}\mathcal{B}w_y)_Y &= -2(p-q, \mathcal{B}w_y)_Y \\ &= -2\langle \mathcal{B}^*(p-q), w_y \rangle = 0 \, , \qquad (12) \end{aligned}$$

since $\mathcal{B}^*(p-q) = -l + l = 0$ in V_0^*.

The 3rd term on the right-hand side of (8) equals $\|\!| \mathcal{A}\mathcal{B}w_y \|\!|_*^2$ that has been estimated from above by $[\![l + \mathcal{B}^*y]\!]^2$ (see (10)). Hence, combining this result with (11) and (12) we obtain from (8):

$$\|\!| \mathcal{B}(u-v) \|\!|^2 + \|\!| p - y \|\!|_*^2 \leq (\|\!| \mathcal{A}\mathcal{B}v - y \|\!|_* + [\![l + \mathcal{B}^*y]\!])^2 + [\![l + \mathcal{B}^*y]\!]^2 , \tag{13}$$

where v is an arbitrary function from $V_0 + u_0$ and y is any function from Y.

From (13) one immediately derives the estimate

$$\left(\|\!| \mathcal{B}(u-v) \|\!|^2 + \|\!| p - y \|\!|_*^2\right)^{1/2} \leq \|\!| \mathcal{A}\mathcal{B}v - y \|\!|_* + \sqrt{2}\, [\![l + \mathcal{B}^*y]\!] \tag{14}$$

and the following theorem:

Theorem 3.1 *Let $(u,p) \in (V_0 + u_0) \times Y$ be the solution of the problem $(\mathcal{M})$, $v \in V_0 + u_0$ and $y \in Y$ arbitrary approximations to u and p.*
Then, the following estimates hold true:

$$\|(u-v, p-y)\|_{V_0 \times Y} \leq \|\!| \mathcal{A}\mathcal{B}v - y \|\!|_* + \sqrt{3}\, [\![l + \mathcal{B}^*y]\!] , \tag{15}$$

$$\|(u-v, p-y)\|_{V_0 \times Y} \geq \frac{1}{\sqrt{3}} (\|\!| \mathcal{A}\mathcal{B}v - y \|\!|_* + [\![l + \mathcal{B}^*y]\!]) . \tag{16}$$

Proof. The upper bound (15) immediately follows from estimate (13) and the definition of the full norm $\|(\cdot,\cdot)\|_{V_0 \times Y}$, since $[\![\mathcal{B}^*(p-y)]\!] = [\![l + \mathcal{B}^*y]\!]$.

To obtain the lower bound (16) we use first the triangle inequality to derive

$$\begin{aligned}\|\!| \mathcal{A}\mathcal{B}v - y \|\!|_* + [\![l + \mathcal{B}^*y]\!] &\leq \|\!| \mathcal{A}\mathcal{B}v - \mathcal{A}\mathcal{B}u \|\!|_* + \|\!| p - y \|\!|_* + [\![l + \mathcal{B}^*y]\!] \\ &= \|\!| \mathcal{B}(u-v) \|\!| + \|\!| p - y \|\!|_* + [\![\mathcal{B}^*(p-y)]\!] ,\end{aligned}$$

and, then, the inequality $a+b+c \leq \sqrt{3}\sqrt{a^2+b^2+c^2}$, $\forall a,b,c \geq 0$, to obtain the estimate

$$\|\!| \mathcal{A}\mathcal{B}v - y \|\!|_* + [\![l + \mathcal{B}^*y]\!] \leq \sqrt{3}\, \|(u-v, p-y)\|_{V_0 \times Y} .$$

This implies the lower bound (16).

■

Let

$$M_\oplus := \|\!| \mathcal{A}\mathcal{B}v - y \|\!|_* + \sqrt{3}\, [\![l + \mathcal{B}^*y]\!] \tag{17}$$

denote the upper bound (15) for the error in the full norm.

Remarks. 1. If $y \to p$ in Y and $v \to u$ in V, the estimates (15) and (16) tend to zero, precisely as the exact error in the full norm $\|(u-v, p-y)\|_{V_0 \times Y}$ does.
2. The error majorant $M_\oplus$ is *sharp*. Indeed, if one takes $y = p$ (i.e. $l + \mathcal{B}^*y \equiv 0$ in V_0^*), estimate (15) becomes

$$\|\!| \mathcal{B}(u-v) \|\!| \leq \|\!| \mathcal{A}\mathcal{B}v - p \|\!|_* = \|\!| \mathcal{B}(u-v) \|\!| ,$$

which shows that the constant "1" in front of the 1st term of $M_\oplus$ cannot be improved, in general. On the other hand, if, in the case $u_0 = 0$, we set $v = 0$ and $y = 0$, then estimate (15) takes the form

$$\left(2\, \|\!| \mathcal{B}u \|\!|^2 + [\![l]\!]^2\right)^{1/2} \leq \sqrt{3} [\![l]\!] ,$$

that is a sharp estimate, since $\|\!| \mathcal{B}u \|\!| \leq [\![l]\!]$ (set $w = u$ in the problem $(\mathcal{P})$) and this estimate, evidently, cannot be improved. Thus, the factor "$\sqrt{3}$" multiplying the 2nd term

in $M_\oplus$ cannot be taken smaller in a general case.
The sharpness of the lower bound (16) in a general case is obvious from the estimate's derivation.
3. The efficiency of the estimator $M_\oplus$ can be easily evaluated using the lower bound (16). Namely, for the effectivity index of $M_\oplus$ we have:

$$\mathrm{i}_{\mathrm{eff}} := \frac{M_\oplus}{\|(u-v,p-y)\|_{V_0\times Y}} \le \sqrt{3}\frac{\| \mathcal{AB}v - y \|_* + \sqrt{3}\,[\![\, l+\mathcal{B}^*y \,]\!]}{\| \mathcal{AB}v - y \|_* + [\![\, l+\mathcal{B}^*y \,]\!]} \le 3\,. \tag{18}$$

Estimate (18) provides a rough upper bound for the effectivity index that in most of the cases will be strictly less than 3. Indeed, if the second term $[\![\, l+\mathcal{B}^*y \,]\!]$ is essentially smaller than the first one, then $\mathrm{i}_{\mathrm{eff}}$ is close to $\sqrt{3}$. On the other hand, since (15) is a guaranteed upper bound of the error, we always have $\mathrm{i}_{\mathrm{eff}} \ge 1$.

The two-sided estimate (15)–(16) is important, because it provides a control over the error in the full norm, i.e. with respect to both primal and dual variables. However, the individual errors in primal and dual variables may also be of interest; in the next two sections we derive sharp upper bounds for the corresponding norms of these errors. It is worth noticing that the individual estimates which immediately follow from (15) are not sharp, hence, may lead to a certain overestimation.

3.2 Estimate for the error in primal variable

Theorem 3.2 *Let $u \in V_0+u_0$ be the solution of the problem $(\mathcal{P})$, $v \in V_0+u_0$ an arbitrary approximate solution to $(\mathcal{P})$.*
Then,

$$\| \mathcal{B}(u-v) \| \le \| \mathcal{AB}v - y \|_* + [\![\, l+\mathcal{B}^*y \,]\!] \quad \forall y \in Y\,. \tag{19}$$

Proof. It immediately follows from (7) that $\| \mathcal{B}(u-v) \|^2 = \inf\limits_{q\in Q_l} \| \mathcal{AB}v - q \|_*^2$, i.e.

$$\| \mathcal{B}(u-v) \| \le \| \mathcal{AB}v - q \|_* \quad \forall q \in Q_l\,. \tag{20}$$

The right-hand side of (20) has been already estimated for the proof of Theorem 3.1, where such function $q \in Q_l$ was constructed that $q = y - \mathcal{AB}w_y$ with y being any function from Y and w_y being the solution to the problem $\mathcal{B}^*\mathcal{AB}w_y = l + \mathcal{B}^*y$ in V_0^*. Then, estimate (19) follows directly from (11).

■

Remarks. 1. Estimate (19) is *sharp*. Indeed, if we set $y = p = \mathcal{AB}u$, the estimate will be

$$\| \mathcal{B}(u-v) \| \le \| \mathcal{AB}v - \mathcal{AB}u \|_* = \| \mathcal{B}(u-v) \| \;.$$

On the other hand, in the case $u_0 = 0$, setting $v = 0$ and $y = 0$ we obtain from (19)

$$\| \mathcal{B}u \| \le [\![\, l \,]\!]$$

that is the sharp energy estimate for the solution u of the problem $(\mathcal{P})$. Thus, the weights equal to 1 on the right-hand side of estimate (19) are optimal, in a general case.
2. Estimate (19) is *asymptotically exact* in the sense that, if $y \to p$ in Y, then the upper bound (19) tends to the norm $\| \mathcal{AB}v - \mathcal{AB}u \|_* = \| \mathcal{B}(u-v) \|$ of the error in primal variable.
3. The estimate remains *efficient*, if y is close to p in Y, since

$$\begin{aligned} \| \mathcal{AB}v - y \|_* + [\![\, l+\mathcal{B}^*y \,]\!] &\le \| \mathcal{AB}v - \mathcal{AB}u \|_* + \| p-y \|_* + [\![\, \mathcal{B}^*(p-y) \,]\!] \\ &\le \| \mathcal{B}(u-v) \| + \sqrt{2}\|p-y\|_{\mathcal{B}^*}\,. \end{aligned}$$

Here the last, presumably small, term measures the level of the overestimation due to estimate (19).
4. If one considers only $y \in Q_l$ in (19), one arrives at estimate (20) that is the "constitutive relation based" estimate (see [19]). On the other hand, if one takes $y = \mathcal{AB}v$ in (19), one obtains the estimate $\|\!| \mathcal{B}(u-v) |\!\| \le [\![l + \mathcal{B}^*\mathcal{AB}v]\!]$ that is the "residual based" estimate for the problem $(\mathcal{P})$ (see [4]). Thus, estimate (19) includes these two estimates as particular cases, combining their advantages and providing a greater flexibility. More on the links between the error majorant and other estimates can be found in [24].

3.3 Estimate for the error in dual variable

Theorem 3.3 *Let $(u,p) \in (V_0 + u_0) \times Y$ be the solution to the problem $(\mathcal{M})$, $y \in Y$ any approximation of p.*
Then,

$$\|\!| p - y |\!\|_* \le \|\!| \mathcal{AB}v - y |\!\|_* + [\![l + \mathcal{B}^* y]\!] \quad \forall v \in V_0 + u_0 \,, \tag{21}$$
$$\|p - y\|_{\mathcal{B}^*} \le \|\!| \mathcal{AB}v - y |\!\|_* + \sqrt{2}[\![l + \mathcal{B}^* y]\!] \quad \forall v \in V_0 + u_0 \,. \tag{22}$$

Proof. We have for any $v \in V_0 + u_0$:

$$\begin{aligned}\|\!| \mathcal{B}(u-v) |\!\|^2 + \|\!| p - y |\!\|_*^2 &= (\mathcal{AB}(u-v), \mathcal{B}(u-v))_Y + (\mathcal{A}^{-1}(p-y), p-y)_Y \\ &= (\mathcal{AB}v - y, \mathcal{B}v - \mathcal{A}^{-1}y)_Y + 2(p-y, \mathcal{B}(u-v))_Y \,, \end{aligned} \tag{23}$$

where the self-adjointness of $\mathcal{A}$ and $\mathcal{A}^{-1}$ as well as the relation $p = \mathcal{AB}u$ have been used. For the 2nd term on the right-hand side of (23) we have:

$$(p-y, \mathcal{B}(u-v))_Y = \langle \mathcal{B}^*(p-y), u-v \rangle = \langle -l - \mathcal{B}^* y, u-v \rangle \quad \forall v \in V_0 + u_0 \,,$$

which implies the estimate

$$|\,(p-y, \mathcal{B}(u-v))_Y\,| \le [\![l + \mathcal{B}^* y]\!] \, \|\!| \mathcal{B}(u-v) |\!\| \le \frac{1}{2}\left([\![l + \mathcal{B}^* y]\!]^2 + \|\!| \mathcal{B}(u-v) |\!\|^2 \right) .$$

Using this estimate and noticing that the 1st term on the right-hand side of (23) equals $\|\!| \mathcal{AB}v - y |\!\|_*^2$, we derive from (23)

$$\|\!| \mathcal{B}(u-v) |\!\|^2 + \|\!| p - y |\!\|_*^2 \le \|\!| \mathcal{AB}v - y |\!\|_*^2 + [\![l + \mathcal{B}^* y]\!]^2 + \|\!| \mathcal{B}(u-v) |\!\|^2 \,,$$

that is

$$\|\!| p - y |\!\|_*^2 \le \|\!| \mathcal{AB}v - y |\!\|_*^2 + [\![l + \mathcal{B}^* y]\!]^2 \quad \forall v \in V_0 + u_0 \,.$$

This immediately yields estimate (21). Then, (22) is obvious.

■

Remark. Estimate (21) is *sharp* (hence, estimate (22) is sharp too). Indeed, if our approximation y belongs to Q_l and we set $v = u$, we obtain from (21)

$$\|\!| p - y |\!\|_* \le \|\!| \mathcal{AB}u - y |\!\|_* = \|\!| p - y |\!\|_* \ .$$

On the other hand, if, in the case $u_0 = 0$, we set $v = 0$ and $y = 0$, we have from (21)

$$\|\!| p |\!\|_* \le [\![l]\!] \,, \text{ i.e. } \|\!| \mathcal{B}u |\!\| \le [\![l]\!] \,,$$

which is the sharp energy estimate for the solution of the problem $(\mathcal{P})$. Thus, the weights of both terms on the right-hand side of (21) are, in general, optimal.

3.4 Important special case

The estimates obtained above provide reliable measures of the errors in a very general situation when the exact solution to the problem $(\mathcal{P})$ is sought in an arbitrary reflexive Banach space V and the given functional l belongs to V_0^*. As a result, the norm in V_0^* enters the estimates, making them less convenient for computational purposes. However, in most of practically interesting cases, one has

$$\text{the } \textit{continuous embedding } \; V \subset U \tag{24}$$

for some Hilbert space U with the inner product $(\cdot,\cdot)_U$ and the norm $\|\cdot\|_U$. Thus, $U \subset V_0^*$ and if, in addition, the given data

$$l \in U\,, \tag{25}$$

we can significantly simplify our estimates.

First, one can notice that, if assumption (25) holds true, the exact dual solution p satisfies the equation $\mathcal{B}^*p + l = 0$ in U and, hence, belongs to the space

$$Y_{\mathcal{B}^*} := \{y \in Y \;\mid\; \mathcal{B}^*y \in U\}$$

that is the Banach space with respect to the norm

$$|\, y \,|_{\mathcal{B}^*} := \big(\| y \|_*^2 + \|\mathcal{B}^*y\|_U^2\big)^{1/2} \quad \forall y \in Y_{\mathcal{B}^*}\,. \tag{26}$$

As compared to the definition of the norm $\|\cdot\|_{\mathcal{B}^*}$ (see (5)), the newly defined norm is stronger, which reflects the fact that $Y_{\mathcal{B}^*}$ is a subspace of Y.

It is natural now to consider the approximation y of the exact dual solution in $Y_{\mathcal{B}^*}$ rather than in Y; this is still much less restrictive than an approximation in the set Q_l whose definition contains the complicated constraint $\mathcal{B}^*y = -l$. Then, we can estimate the term $[\![\, l + \mathcal{B}^*y \,]\!]$ as follows:

$$\begin{aligned}
[\![\, l{+}\mathcal{B}^*y \,]\!] &= \sup_{w\in V_0\setminus\{0\}} \frac{\langle l + \mathcal{B}^*y, w\rangle}{\| \mathcal{B}w \|} = \sup_{w\in V_0\setminus\{0\}} \frac{(l + \mathcal{B}^*y, w)_U}{\| \mathcal{B}w \|} \le \sup_{w\in V_0\setminus\{0\}} \frac{\|l + \mathcal{B}^*y\|_U \, \|w\|_U}{\| \mathcal{B}w \|} \\
&\le \sup_{w\in V_0\setminus\{0\}} \frac{\|l + \mathcal{B}^*y\|_U \, \|w\|_V}{\| \mathcal{B}w \|} \le \frac{C_{\mathcal{B}}}{\lambda_A^{1/2}} \|l + \mathcal{B}^*y\|_U \quad \forall y \in Y_{\mathcal{B}^*}\,,
\end{aligned} \tag{27}$$

where the continuity of the embedding $V \subset U$ as well as inequalities (3) and (4) have been used. It is important to notice that one often has the inequality

$$\|w\|_U \;\le\; \widetilde{C}_{\mathcal{B}}\, \|\mathcal{B}w\|_Y \quad \forall w \in V_0 \tag{28}$$

in addition to (4); in such a case, the constant $C_{\mathcal{B}}$ in (27) is to be replaced by $\widetilde{C}_{\mathcal{B}}$ from (28).

With the definition of the $Y_{\mathcal{B}^*}$-norm (see (26)), the *full norm* (6) should be understood on the product space $V_0 \times Y_{\mathcal{B}^*}$ in the following sense:

$$\|(v,y)\|_{V_0\times Y_{\mathcal{B}^*}} := \big(\| \mathcal{B}v \|^2 + \| y \|_*^2 + \|\mathcal{B}^*y\|_U^2\big)^{1/2} \quad \forall (v,y) \in V_0 \times Y_{\mathcal{B}^*}\,. \tag{29}$$

Theorem 3.4 *Let V be continuously embedded into some Hilbert space U and $l \in U$. Let $(u,p) \in (V_0 + u_0) \times Y_{\mathcal{B}^*}$ be the solution to the problem $(\mathcal{M})$ and $(v,y) \in (V_0 + u_0) \times Y_{\mathcal{B}^*}$*

any approximate solution to $(\mathcal{M})$.
Then, the following a posteriori error estimates hold true:

$$\|(u-v,p-y)\|_{V_0\times Y_{\mathcal{B}^*}} \leq ⦀ \mathcal{A}\mathcal{B}v - y ⦀_* + \left(1+2\frac{C_{\mathcal{B}}^2}{\lambda_A}\right)^{1/2} \|l+\mathcal{B}^*y\|_U \,, \tag{30}$$

$$\|(u-v,p-y)\|_{V_0\times Y_{\mathcal{B}^*}} \geq \frac{1}{\sqrt{3}}\left(⦀ \mathcal{A}\mathcal{B}v - y ⦀_* + \|l+\mathcal{B}^*y\|_U\right) \,, \tag{31}$$

$$⦀ \mathcal{B}(u-v) ⦀ \leq ⦀ \mathcal{A}\mathcal{B}v - y ⦀_* + \frac{C_{\mathcal{B}}}{\lambda_A^{1/2}} \|l+\mathcal{B}^*y\|_U \,, \tag{32}$$

$$⦀ p-y ⦀_* \leq ⦀ \mathcal{A}\mathcal{B}v - y ⦀_* + \frac{C_{\mathcal{B}}}{\lambda_A^{1/2}} \|l+\mathcal{B}^*y\|_U \,, \tag{33}$$

$$\|p-y\|_{\mathcal{B}^*} \leq ⦀ \mathcal{A}\mathcal{B}v - y ⦀_* + \left(1+\frac{C_{\mathcal{B}}^2}{\lambda_A}\right)^{1/2} \|l+\mathcal{B}^*y\|_U \,. \tag{34}$$

Proof. The upper bound (30) immediately follows from estimates (13) and (27); the lower bound (31) is a simple consequence of the triangle inequality, like the lower bound (16) in Theorem 3.1.

Estimate (32) can be easily derived from (19) and (27), while estimates (33) and (34) follow from (21) and (27).

■

Remarks. 1. Estimates (30)–(34) are *sharp*, which follows from the sharpness of the estimates of Theorems 3.1–3.3 and of inequality (27).
2. Estimates (30) and (31) imply that the effectivity index of the error majorant (30) is always between 1 and $\sqrt{3}\left(1+2\frac{C_{\mathcal{B}}^2}{\lambda_A}\right)^{1/2}$. It is worth noting that the constant λ_A can be made equal to 1, if one performs the corresponding rescaling of the operator $\mathcal{A}$ and of the functional l (i.e. the multiplication of the linear problem $(\mathcal{P})$ by $1/\lambda_A$). The constant $C_{\mathcal{B}}$ depends only on the operator $\mathcal{B}$ and can be easily evaluated a priori (we discuss this issue in the next section). We will also show that, after an appropriate scaling of the geometric coordinates, one can make the constant $C_{\mathcal{B}} \leq 1$, which means that the effectivity index of the upper bound (30) for the new, "rescaled" problem will always be between 1 and 3.
3. It is worthwhile to notice a remarkable symmetry of estimates (32) and (33) for the primal and dual variables.

4 Applications

4.1 Diffusion problem

4.1.1 Error estimates

Let $V = H^1(\Omega)$, where Ω is a bounded domain in $\mathbb{R}^n$ with Lipschitz boundary $\partial\Omega$, $V_0 = H_0^1(\Omega)$, $Y = L_2(\Omega;\mathbb{R}^n)$. Consider the case

$$\mathcal{B} = \nabla := \left(\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_n}\right) .$$

Then, $\mathcal{B}^*\mathbf{y} = -\operatorname{div}\mathbf{y} \in H^{-1}(\Omega) = V_0^*$ for any $\mathbf{y} \in Y$, and

$$\langle \mathcal{B}^*\mathbf{y}, w\rangle = \int_\Omega \mathbf{y}\cdot\nabla w\, dx \quad \forall w \in V_0 \,,$$

where the dot denotes the scalar product of vectors in $\mathbb{R}^n$. The operator $\mathcal{A}$ is defined by a symmetric, uniformly positive definite matrix $\mathbf{A} = \{a_{ij}(x)\}_{i,j=\overline{1,n}}$ with coefficients from $L_\infty(\Omega)$. Then, the norms $|\!|\!| \cdot |\!|\!|$ and $|\!|\!| \cdot |\!|\!|_*$ are defined as

$$|\!|\!| \mathbf{y} |\!|\!|^2 = \int_\Omega \mathbf{A}\mathbf{y} \cdot \mathbf{y}\, dx\,, \quad |\!|\!| \mathbf{y} |\!|\!|_*^2 = \int_\Omega \mathbf{A}^{-1}\mathbf{y} \cdot \mathbf{y}\, dx\,.$$

Inequality (3) is obviously satisfied, and (4) follows from the Friedrichs inequality.

Assume now that u_0 is some given function from $H^1(\Omega)$ and l is some given functional from $H^{-1}(\Omega)$. Then, the problem $(\mathcal{P})$ defines the weak solution of the boundary-value problem

$$-\mathrm{div}\,(\mathbf{A}\nabla u) + l = 0 \quad \text{in}\ \Omega\,, \tag{35}$$

$$u = u_0 \quad \text{on}\ \partial\Omega\,. \tag{36}$$

We can also write all the estimates of Theorems 3.1–3.3, where $[\![\cdot]\!]$ is equivalent to the $H^{-1}(\Omega)$-norm.

We see, however, that V is continuously embedded into the Hilbert space $U = L_2(\Omega)$, hence, if we suppose that the data $l \in L_2(\Omega)$, we can use the results of Theorem 3.4.

First, we note that the space $Y_{\mathcal{B}^*}$ is, in fact, the space $H(\Omega; \mathrm{div}) := \{\mathbf{y} \in L_2(\Omega; \mathbb{R}^n) \mid \mathrm{div}\,\mathbf{y} \in L_2(\Omega)\}$ with the norm

$$\|\mathbf{y}\|_{\mathrm{div}} := \left(|\!|\!| \mathbf{y} |\!|\!|_*^2 + \|\mathrm{div}\,\mathbf{y}\|^2\right)^{1/2}\,,$$

where $\|\cdot\|$ denotes the norm in $L_2(\Omega)$.

The *full norm* takes, then, the form

$$\|(v, \mathbf{y})\|_{1\times\mathrm{div}} := \left(|\!|\!| \nabla v |\!|\!|^2 + |\!|\!| \mathbf{y} |\!|\!|_*^2 + \|\mathrm{div}\,\mathbf{y}\|^2\right)^{1/2} \quad \forall (v, \mathbf{y}) \in H_0^1(\Omega) \times H(\Omega; \mathrm{div})\,.$$

It is important to notice that, in the considered case, we have the inequality of type (28) that is exactly the Friedrichs inequality

$$\|w\| \leq C_\Omega \|\nabla w\| \quad \forall w \in H_0^1(\Omega)\,.$$

Thus, the constant $C_{\mathcal{B}}$ in (27) and in Theorem 3.4 is, in fact, the constant C_Ω from the Friedrichs inequality.

Hence, if $(u, \mathbf{p}) \in (H_0^1(\Omega) + u_0) \times H(\Omega; \mathrm{div})$ is the exact solution to the mixed problem

$$\mathbf{p} = \mathbf{A}\nabla u \quad \text{in}\ \Omega\,, \tag{37}$$

$$-\mathrm{div}\,\mathbf{p} + l = 0 \quad \text{in}\ \Omega\,, \tag{38}$$

and $(v, \mathbf{y}) \in (H_0^1(\Omega) + u_0) \times H(\Omega; \mathrm{div})$ is any approximate solution to the problem, then the following a posteriori error estimates follow directly from Theorem 3.4:

$$\|(u - v, \mathbf{p} - \mathbf{y})\|_{1\times\mathrm{div}} \leq |\!|\!| \mathbf{A}\nabla v - \mathbf{y} |\!|\!|_* + \left(1 + 2\frac{C_\Omega^2}{\lambda_A}\right)^{1/2} \|\mathrm{div}\,\mathbf{y} - l\|\,, \tag{39}$$

$$\|(u - v, \mathbf{p} - \mathbf{y})\|_{1\times\mathrm{div}} \geq \frac{1}{\sqrt{3}} \left(|\!|\!| \mathbf{A}\nabla v - \mathbf{y} |\!|\!|_* + \|\mathrm{div}\,\mathbf{y} - l\|\right)\,, \tag{40}$$

$$|\!|\!| \nabla(u - v) |\!|\!| \leq |\!|\!| \mathbf{A}\nabla v - \mathbf{y} |\!|\!|_* + \frac{C_\Omega}{\lambda_A^{1/2}} \|\mathrm{div}\,\mathbf{y} - l\|\,, \tag{41}$$

$$|\!|\!| \mathbf{p} - \mathbf{y} |\!|\!|_* \leq |\!|\!| \mathbf{A}\nabla v - \mathbf{y} |\!|\!|_* + \frac{C_\Omega}{\lambda_A^{1/2}} \|\mathrm{div}\,\mathbf{y} - l\|\,, \tag{42}$$

$$\|\mathbf{p} - \mathbf{y}\|_{\mathrm{div}} \leq |\!|\!| \mathbf{A}\nabla v - \mathbf{y} |\!|\!|_* + \left(1 + \frac{C_\Omega^2}{\lambda_A}\right)^{1/2} \|\mathrm{div}\,\mathbf{y} - l\|\,. \tag{43}$$

Estimates (39)–(43) provide sharp error bounds that are explicitly computable, if one has the approximate solution to (37)–(38) in the product space $H^1(\Omega) \times H(\Omega; \mathrm{div})$. It is, of course, clear that having found the approximate mixed solution $(v, \mathbf{y})$ by primal or dual mixed FEM one can use some local averaging (projection) to recover the needed $H^1(\Omega)$, respectively, $H(\Omega; \mathrm{div})$ regularity for the approximate primal, respectively, dual variable. There exist, however, several methods allowing to approximate the mixed solution $(u, \mathbf{p})$ in the space $H^1(\Omega) \times H(\Omega; \mathrm{div})$ directly. Below, we briefly review four of them.

4.1.2 Approximation of the mixed solution in $H^1(\Omega) \times H(\Omega; \mathrm{div})$

a) Least-squares mixed method
The method was analysed in [21] and, under the name First-Order-System Least-Squares (FOSLS), in [10], [11] (see also the references therein). In this method, the saddle-point (min-max) problem (37)–(38) is reformulated as a quadratic minimization (min-min) problem

$$\inf_{v \in H_0^1(\Omega)+u_0} \; \inf_{\mathbf{y} \in H(\Omega;\mathrm{div})} \left(\| \mathbf{A}\nabla v - \mathbf{y} \|_*^2 + \|\mathrm{div}\, \mathbf{y} - l\|^2 \right) , \tag{44}$$

which leads to the solution of the "stabilized" saddle-point problem:
Find $(u, \mathbf{p}) \in (H_0^1(\Omega) + u_0) \times H(\Omega; \mathrm{div})$ such that

$$\int_\Omega \mathbf{A}^{-1}\mathbf{p} \cdot \mathbf{q}\, dx + \int_\Omega (\mathrm{div}\, \mathbf{p})(\mathrm{div}\, \mathbf{q})\, dx - \int_\Omega \nabla u \cdot \mathbf{q}\, dx = \int_\Omega l\, (\mathrm{div}\, \mathbf{q})\, dx \quad \forall \mathbf{q} \in H(\Omega; \mathrm{div})\,, \tag{45}$$

$$\int_\Omega (\mathrm{div}\, \mathbf{p})\, v\, dx + \int_\Omega \mathbf{A}\nabla u \cdot \nabla v\, dx = 0 \quad \forall v \in H_0^1(\Omega)\,. \tag{46}$$

We have to note that, in the original version of the method, the squared L_2-norm was used in the 1st term of the functional (44) instead of the squared $\| \cdot \|_*$-norm, which somehow changes the system of the functional's optimality conditions (45)–(46).

System (45)–(46), unlike (37)–(38), leads to symmetric *positive definite* discrete problem, and the discrete *inf-sup* condition is always satisfied owing to the least-squares stabilization. The latter fact allows one to choose the approximation spaces for u and $\mathbf{p}$ independently of each other.

However, in (45)–(46) the primal and the dual variables are *strongly coupled.* The following method yields only a weak coupling of the variables.

b) Method of minimizing the squared majorant
From estimate (41) for the error in the primal variable one can easily derive the estimate for the squared energy norm of the error:

$$\| \nabla(u - v) \|^2 \leq (1 + \beta)\, \| \mathbf{A}\nabla v - \mathbf{y} \|_*^2 + \left(1 + \frac{1}{\beta}\right) \frac{C_\Omega^2}{\lambda_A} \|\mathrm{div}\, \mathbf{y} - l\|^2\,, \tag{47}$$

where $\beta > 0$ is an arbitrary number and $\mathbf{y}$ is any function from $H(\Omega; \mathrm{div})$. Denote the right-hand side of (47) by $M^2(v; \mathbf{y}, \beta)$ ("the squared error majorant"). It is evident that $M^2(v; \mathbf{y}, \beta)$ is, in fact, the least-squares functional (44) with differently weighted terms. However, instead of minimizing the functional with respect to both v and $\mathbf{y}$ simultaneously as in the least-squares mixed method, the following simple algorithm was proposed in [26]:

1) Find the approximate solution $v \in V_0 + u_0$ to the problem (35)–(36).

2) Set $\beta = 1$ and find $\mathbf{y}$ by minimizing $M^2(v; \mathbf{y}, \beta)$ with respect to $\mathbf{y}$.

The algorithm was initially motivated by the goal to find a best possible upper bound for the energy error in the primal variable; however, it also provides a computationally efficient way of computing approximate primal and dual solutions in $H^1(\Omega) \times H(\Omega; \mathrm{div})$ in a *weakly coupled* manner. Indeed, the problems for v and $\mathbf{y}$ have now to be solved successively.

While the problem of finding an approximate solution to (35)–(36) on step 1 is quite standard, the computation of $\mathbf{y}$ on step 2 does not present serious difficulty as well. Since $M^2(v; \mathbf{y}, \beta)$ is a quadratic functional with respect to the dual variable $\mathbf{y}$ for any fixed v and β, the minimization of the functional on any finite-dimensional subspace Y_h of $H(\Omega; \mathrm{div})$ leads to the solution of a linear system with symmetric positive definite matrix.

This algorithm has been independently proposed in [8] as an alternative to the least-squares mixed method and considered as a single Picard-Uzawa type iteration for the solution of the coupled system (45)–(46) (in [8], the least-squares functional (44) was used, not $M^2(v; \mathbf{y}, \beta)$). It has been shown in [8] that, with v found by a conforming FEM for the problem (35)–(36), the minimizer $\mathbf{y}_h$ of the functional on the subspace Y_h has the optimal order of the $H(\Omega; \mathrm{div})$-error with respect to the mesh size (i.e. the order of the interpolation error for Y_h), provided that the $H^1(\Omega)$-error of the approximation v is not of lower order. The advantage over the dual mixed FEM as well as the least-squares mixed method is obvious: the computation of the primal variable is completely independent of the calculation of the dual one (this reduces the total computational cost), and the discrete problem for each of the variables is moderately-sized, symmetric and positive definite (i.e. one does not have to deal with an indefinite saddle-point problem as in the case of the dual mixed FEM).

As follows from the numerical studies of [26], using the parameter β one can gain a further improvement in the approximation of the dual solution. Namely, for the unique minimizer $\mathbf{y}_\beta \in H(\Omega; \mathrm{div})$ of $M^2(v; \mathbf{y}, \beta)$ for any fixed $v \in V_0 + u_0$ and $\beta > 0$, it was proved in [26] that $\mathbf{y}_\beta$ converges to the exact dual solution p in $H(\Omega; \mathrm{div})$ as $\beta \to 0$, and, moreover,

$$
\begin{aligned}
|\!|\!| \mathbf{p} - \mathbf{y} |\!|\!|_* &\le C\,\beta^{1/2}\,, \\
\|\mathrm{div}\,(\mathbf{p} - \mathbf{y})\| &\le C\,\beta\,,
\end{aligned}
$$

with some constant C independent of $\mathbf{y}$ and β. Thus, the one-stroke minimization of the functional $M^2(v; \mathbf{y}, \beta)$ with respect to $\mathbf{y}$ and with some moderately small β may yield even better accuracy of the dual-solution approximation than the minimization with $\beta = 1$. Numerical experiments (see [26]) show that, for example, if one uses linear finite elements for both primal and dual variables, the value $\beta = 1/10$ is a good choice. Taking β moderately small allows to circumvent the difficulties with the condition number of the resulting discrete system and with the locking phenomenon, typical for the penalty methods.

A possible way of finding the concrete value of β is to minimize the functional $M^2(v; \mathbf{y}, \beta)$ with respect to β having fixed v and $\mathbf{y}$. This immediately implies the explicit formula

$$
\beta = \frac{C_\Omega \|\mathrm{div}\,\mathbf{y} - l\|}{\lambda_A^{1/2}\ |\!|\!| \mathbf{A}\nabla v - \mathbf{y} |\!|\!|_*}\,, \tag{48}
$$

and the modified algorithm for the approximation of the primal and dual solutions reads:

1) Find the approximate solution $v \in V_0 + u_0$ to the problem (35)–(36).

2) Set $\beta^{(1)} = 1$ and find $\mathbf{y}^{(1)}$ by minimizing $M^2(v; \mathbf{y}, \beta^{(1)})$ with respect to $\mathbf{y}$.

3) Compute $\beta^{(2)}$ using $\mathbf{y}^{(1)}$ in formula (48); find $\mathbf{y}^{(2)}$ by minimizing $M^2(v;\mathbf{y},\beta^{(2)})$ with respect to $\mathbf{y}$.

A further iteration of the process of minimizing the squared majorant with respect to $\mathbf{y}$ and β does not bring any essential benefits, as shown in the detailed study of [26].

To summarize, the minimization of the squared majorant either at a one stroke (only the steps 1 and 2 in the algorithm above) or by two iterations provides a competitive approach to the approximation of the dual solution. The whole method of finding the primal variable in $H^1(\Omega)$ and the dual variable in $H(\Omega;\mathrm{div})$ amounts to the successive solution of two elliptic problems. It is worth noting that the approximation spaces for v and $\mathbf{y}$ can be chosen independently of each other, as in the least-squares mixed method.

c) Dual penalty method

This method can be viewed as a limiting case of the previous method, i.e. the case when the parameter β in the squared majorant is considered as a very small penalty parameter. The classical dual penalty method has, however, a slightly different formulation. Namely, after finding $v \in V_0 + u_0$ as an approximate solution to (35)–(36), one has to minimize the quadratic "penalized functional"

$$I(\mathbf{y}) := \| \mathbf{y} \|_*^2 + \frac{1}{\varepsilon}\|\mathrm{div}\,\mathbf{y} - l\|^2$$

over $H(\Omega;\mathrm{div})$ for some small $\varepsilon > 0$. The main difference with the method of minimizing the squared majorant is that the approximation of $\mathbf{y}$ is now *fully decoupled* from the approximation of v. As immediate drawbacks one has the deterioration of the condition number of the resulting discrete problem and possible locking phenomenon.

d) Method of local projections

In this method, the dual variable is found by some local projections of the approximate flux $\mathbf{A}\nabla v$ into the space $H(\Omega;\mathrm{div})$. The approximate flux is derived from the approximate primal solution $v \in V_0 + u_0$ previuosly found by solving (35)–(36). Thus, we have here again a *weakly coupled* approach. The method is usually referred to as the "gradient recovery" or "gradient averaging", and its diverse variants have been considered by many researchers (see, e.g., [18], [33], [34], [32] and the references therein). In particular, the so-called "equilibrium-enhanced" gradient recovery methods (see [5], [30]) seem to be especially advantageous for computing an accurate approximation to the dual variable in the $H(\Omega;\mathrm{div})$-norm.

Remarks. 1. It is clear that each of the four methods addressed above has both advantages and drawbacks. The thorough comparison of the methods still remains to be done.
2. If the approximation to $(u,\mathbf{p})$ has been found in $(H_0^1(\Omega)+u_0)\times H(\Omega;\mathrm{div})$ by one of the above considered methods, it can be inserted into estimates (39)–(43) to yield the explicit a posteriori control of the errors in both variables. Since the norms in the estimates can be computed by summation of the local contributions from subdomains of Ω (given some finite subdivision of Ω), they may be used also for an *adaptive* improvement of the approximation. In particular, it is obvious that, if $\mathbf{y}$ is close to $\mathbf{p}$ in $H(\Omega;\mathrm{div})$, the term $\| \mathbf{A}\nabla v - \mathbf{y} \|_*$ computed over any subdomain $\omega \subset \Omega$ is close to $\| \nabla(u-v) \|$ considered on ω. More on the use of the error majorants for the indication of the local error distribution can be found in [26], [27].
3. The constant C_Ω stemming from Friedrichs' inequality is equal to $1/\sqrt{\lambda_\Omega}$, where λ_Ω is the minimal eigenvalue of the Laplace operator equipped with the homogeneous Dirichlet boundary condition on $\partial\Omega$. It is, however, clear that C_Ω can always be estimated from above by $C_{\mathcal{D}}$, where $\mathcal{D} \supset \Omega$ is some domain of a simple shape (e.g., a rectangle in 2D).

Then, $C_{\mathcal{D}}$ can be computed analytically.
4. Since the evaluation of C_Ω is fairly easy, the total computational cost of estimates (39)–(43) is very small (only the computation of norms), provided the pair $(v, \mathbf{y})$ is found in $(H_0^1(\Omega) + u_0) \times H(\Omega; \mathrm{div})$, for instance, by one of the methods discussed above.
5. Using translation and rescaling of the geometric coordinates (which amounts to a linear coordinate transformation), one can make it so that the rescaled physical domain $\widetilde{\Omega}$ would be completely inside of a unit cube (square in 2D). After having rewritten the original elliptic problem in the new coordinates and subsequent rescaling the equation so that $\lambda_A = 1$ (see Remark 2 at the end of section 3.4), we can write down all the estimates (39)–(43) for the approximation error of the solution to the rescaled problem on the new domain $\widetilde{\Omega}$. The most important fact here is that all the estimates for the new problem will contain only numerical constants (like $\sqrt{3}$, $\sqrt{2}$), since $\lambda_A = 1$ and the Friedrichs constant $C_{\widetilde{\Omega}}$ may be estimated from above by 1 (see Remark 3 above). As immediate consequence, one infers that the effectivity index of the upper bound (39) for the error in the full norm will be between 1 and 3.

4.2 Linear elasticity

Although the application of the theory to the problem of linear elasticity is similar to the case of the diffusion problem, it is, however, interesting to consider the elasticity problem in detail.

Let $V = H^1(\Omega; \mathbb{R}^n)$, where Ω is a bounded domain in $\mathbb{R}^n$ with Lipschitz boundary $\partial\Omega$, $V_0 = \{\mathbf{v} \in V \mid \mathbf{v} = 0 \text{ on } \partial\Omega\}$, and $\mathbf{Y} = L_2(\Omega; \mathbb{M}_s^{n\times n})$, where $\mathbb{M}_s^{n\times n}$ is the space of symmetric $n \times n$-matrices. Define now the operator $\mathcal{B}$ as follows:

$$\mathcal{B}\mathbf{v} := \underline{\underline{\mathbf{e}}}(\mathbf{v}) = \frac{1}{2}\left(\nabla\mathbf{v} + (\nabla\mathbf{v})^T\right) .$$

Here $\nabla\mathbf{v} = \{v_{i,j}\}$ is a tensor (the gradient of the vector $\mathbf{v}$) and the symbol ^{T} means the transposition. Then, $\mathcal{B}^*\underline{\underline{\mathbf{y}}} = -\mathrm{div}\,\underline{\underline{\mathbf{y}}} \in H^{-1}(\Omega; \mathbb{R}^n) = V_0^*$ for any $\underline{\underline{\mathbf{y}}} \in \mathbf{Y}$, and

$$\langle \mathcal{B}^*\underline{\underline{\mathbf{y}}}, \mathbf{w}\rangle = \int_\Omega \underline{\underline{\mathbf{y}}} : \underline{\underline{\mathbf{e}}}(\mathbf{w})\, dx \quad \forall \mathbf{w} \in V_0 ,$$

where the colon denotes the inner product in $\mathbb{M}_s^{n\times n}$ ($a\colon b = \sum a_{ij}b_{ij}\ \forall a, b \in \mathbb{M}_s^{n\times n}$).

The operator $\mathcal{A}$ is defined by the so-called tensor of elastic moduli $\mathbb{L} = \{\mathbb{L}_{ijkl}\}$ that satisfies the double inequality

$$\lambda_{\mathbb{L}} \mid \underline{\underline{\mathbf{e}}} \mid^2 \le \mathbb{L}\underline{\underline{\mathbf{e}}} : \underline{\underline{\mathbf{e}}} \le \Lambda_{\mathbb{L}} \mid \underline{\underline{\mathbf{e}}} \mid^2 \quad \forall \underline{\underline{\mathbf{e}}} \in \mathbb{M}_s^{n\times n} , \tag{49}$$

and the symmetry and boundedness conditions

$$\mathbb{L}_{ijkl} = \mathbb{L}_{jikl} = \mathbb{L}_{klij} , \quad \mathbb{L}_{ijkl} \in L_\infty(\Omega) . \tag{50}$$

Then, the norms $|\!|\!| \cdot |\!|\!|$ and $|\!|\!| \cdot |\!|\!|_*$ are defined as

$$|\!|\!| \underline{\underline{\mathbf{y}}} |\!|\!|^2 = \int_\Omega \mathbb{L}\underline{\underline{\mathbf{y}}} : \underline{\underline{\mathbf{y}}}\, dx , \quad |\!|\!| \underline{\underline{\mathbf{y}}} |\!|\!|_*^2 = \int_\Omega \mathbb{L}^{-1}\underline{\underline{\mathbf{y}}} : \underline{\underline{\mathbf{y}}}\, dx \quad \forall \underline{\underline{\mathbf{y}}} \in \mathbf{Y} .$$

Inequality (3) is obviously satisfied, and (4) follows from the Korn inequality.

Assume now that $\mathbf{u}_0$ is some given function from $H^1(\Omega; \mathbb{R}^n)$ and $\mathbf{f}$ is some given functional from $H^{-1}(\Omega; \mathbb{R}^n)$. Then, the problem $(\mathcal{P})$ can be formulated as follows:
Find $\mathbf{u} \in V_0 + \mathbf{u}_0$ such that

$$\int_\Omega \mathbb{L}\underline{\underline{\mathbf{e}}}(\mathbf{u}) : \underline{\underline{\mathbf{e}}}(\mathbf{w})\, dx + \langle \mathbf{f}, \mathbf{w}\rangle = 0 \quad \forall \mathbf{w} \in V_0 , \tag{51}$$

where $\langle \mathbf{f}, \mathbf{w} \rangle = \int_\Omega \mathbf{f} \cdot \mathbf{w}\, dx$. The corresponding mixed formulation of (51) defines the weak solution of the boundary-value problem of linear elasticity:

$$\underline{\underline{\mathbf{p}}} = \mathbb{L}\underline{\underline{\mathbf{e}}}(\mathbf{u}) \quad \text{in } \Omega\,, \tag{52}$$

$$\operatorname{div} \underline{\underline{\mathbf{p}}} = \mathbf{f} \quad \text{in } \Omega\,, \tag{53}$$

$$\mathbf{u} = \mathbf{u}_0 \quad \text{on } \partial\Omega\,. \tag{54}$$

We see that V is continuously embedded into the Hilbert space $U = L_2(\Omega;\mathbb{R}^n)$, hence, if we suppose that the given body force $\mathbf{f} \in L_2(\Omega;\mathbb{R}^n)$, we can use the results of Theorem 3.4.

First, we note that the space $Y_{\mathcal{B}^*}$ is, in fact, the space $\mathbf{H}(\Omega;\operatorname{div}) := \{\underline{\underline{\mathbf{y}}} \in \mathbf{Y} \mid \operatorname{div}\underline{\underline{\mathbf{y}}} \in L_2(\Omega;\mathbb{R}^n)\}$ with the norm

$$\|\underline{\underline{\mathbf{y}}}\|_{\operatorname{div}} := \left(|\!|\!| \underline{\underline{\mathbf{y}}} |\!|\!|_*^2 + \|\operatorname{div}\underline{\underline{\mathbf{y}}}\|^2 \right)^{1/2},$$

where $\|\cdot\|$ denotes the norm in $L_2(\Omega;\mathbb{R}^n)$.

The *full norm* takes, then, the form

$$\|(\mathbf{v},\underline{\underline{\mathbf{y}}})\|_{1\times\operatorname{div}} := \left(|\!|\!| \underline{\underline{\mathbf{e}}}(\mathbf{v}) |\!|\!|^2 + |\!|\!| \underline{\underline{\mathbf{y}}} |\!|\!|_*^2 + \|\operatorname{div}\underline{\underline{\mathbf{y}}}\|^2 \right)^{1/2} \quad \forall (\mathbf{v},\underline{\underline{\mathbf{y}}}) \in V_0 \times \mathbf{H}(\Omega;\operatorname{div})\,.$$

It is important to notice that, in the considered case, we have the inequality of type (28) that is a vector variant of the Friedrichs inequality, namely,

$$\|\mathbf{w}\| \le C_\Omega \,\|\underline{\underline{\mathbf{e}}}(\mathbf{w})\| \quad \forall \mathbf{w} \in V_0\,. \tag{55}$$

Thus, the constant $C_{\mathcal{B}}$ in (27) and in Theorem 3.4 is, in fact, the constant C_Ω from (55).

Remark. The constant C_Ω from (55) equals $1/\sqrt{\lambda_\Omega}$, where λ_Ω is the minimal eigenvalue of the vector-valued elliptic operator $\mathcal{L} : V_0 \to H^{-1}(\Omega;\mathbb{R}^n)$, $\mathcal{L}\mathbf{w} = -\frac{1}{2}\left(\operatorname{div}(\nabla\mathbf{w}) + \nabla(\operatorname{div}\mathbf{w})\right)$ for any $\mathbf{w} \in V_0$, equipped with the zero Dirichlet boundary condition on $\partial\Omega$. The minimal eigenvalue λ_Ω can be estimated from below by the one half of the sum of the minimal eigenvalues of the operators $\mathcal{L}_1 : V_0 \to H^{-1}(\Omega;\mathbb{R}^n)$, $\mathcal{L}_1\mathbf{w} = -\operatorname{div}(\nabla\mathbf{w}) = -\Delta\mathbf{w}$, and $\mathcal{L}_2 : V_0 \to H^{-1}(\Omega;\mathbb{R}^n)$, $\mathcal{L}_2\mathbf{w} = -\nabla(\operatorname{div}\mathbf{w})$. It is clear that the smallest eigenvalue of the 2nd operator is zero, while the minimal eigenvalue of the 1st one equals the minimal eigenvalue of the scalar Laplace operator in Ω; the latter depends only on the geometry of the domain Ω and can be estimated from below by embedding Ω into a larger domain of a simpler shape, as discussed in Remark 3 at the end of section 4.1. This ultimately leads to an easily computable upper bound for the constant C_Ω from (55).

Hence, if $(\mathbf{u},\underline{\underline{\mathbf{p}}}) \in (V_0 + \mathbf{u}_0) \times \mathbf{H}(\Omega;\operatorname{div})$ is the exact solution to the mixed problem (52)–(54) and $(\mathbf{v},\underline{\underline{\mathbf{y}}}) \in (V_0 + \mathbf{u}_0) \times \mathbf{H}(\Omega;\operatorname{div})$ is any approximate solution to the problem, then the following a posteriori error estimates follow directly from Theorem 3.4:

$$\|(\mathbf{u}-\mathbf{v},\underline{\underline{\mathbf{p}}}-\underline{\underline{\mathbf{y}}})\|_{1\times\operatorname{div}} \le |\!|\!| \mathbb{L}\underline{\underline{\mathbf{e}}}(\mathbf{v}) - \underline{\underline{\mathbf{y}}} |\!|\!|_* + \left(1 + 2\frac{C_\Omega^2}{\lambda_{\mathbb{L}}}\right)^{1/2} \|\operatorname{div}\underline{\underline{\mathbf{y}}} - \mathbf{f}\|\,, \tag{56}$$

$$\|(\mathbf{u}-\mathbf{v},\underline{\underline{\mathbf{p}}}-\underline{\underline{\mathbf{y}}})\|_{1\times\operatorname{div}} \ge \frac{1}{\sqrt{3}} \left(|\!|\!| \mathbb{L}\underline{\underline{\mathbf{e}}}(\mathbf{v}) - \underline{\underline{\mathbf{y}}} |\!|\!|_* + \|\operatorname{div}\underline{\underline{\mathbf{y}}} - \mathbf{f}\| \right), \tag{57}$$

$$|\!|\!| \underline{\underline{\mathbf{e}}}(\mathbf{u}-\mathbf{v}) |\!|\!| \le |\!|\!| \mathbb{L}\underline{\underline{\mathbf{e}}}(\mathbf{v}) - \underline{\underline{\mathbf{y}}} |\!|\!|_* + \frac{C_\Omega}{\lambda_{\mathbb{L}}^{1/2}} \|\operatorname{div}\underline{\underline{\mathbf{y}}} - \mathbf{f}\|\,, \tag{58}$$

$$|\!|\!| \underline{\underline{\mathbf{p}}} - \underline{\underline{\mathbf{y}}} |\!|\!|_* \le |\!|\!| \mathbb{L}\underline{\underline{\mathbf{e}}}(\mathbf{v}) - \underline{\underline{\mathbf{y}}} |\!|\!|_* + \frac{C_\Omega}{\lambda_{\mathbb{L}}^{1/2}} \|\operatorname{div}\underline{\underline{\mathbf{y}}} - \mathbf{f}\|\,, \tag{59}$$

$$\|\underline{\underline{\mathbf{p}}} - \underline{\underline{\mathbf{y}}}\|_{\operatorname{div}} \le |\!|\!| \mathbb{L}\underline{\underline{\mathbf{e}}}(\mathbf{v}) - \underline{\underline{\mathbf{y}}} |\!|\!|_* + \left(1 + \frac{C_\Omega^2}{\lambda_{\mathbb{L}}}\right)^{1/2} \|\operatorname{div}\underline{\underline{\mathbf{y}}} - \mathbf{f}\|\,. \tag{60}$$

Estimates (56)–(60) provide sharp error bounds that are explicitly computable, if one has the approximate solution to problem (52)–(54) in the product space $(V_0 + \mathbf{u}_0) \times \mathbf{H}(\Omega; \mathrm{div})$. The construction of the approximation in this space can be done along the lines presented in the previous section for the case of a scalar elliptic problem.

References

[1] Y. Achdou, C. Bernardi and F. Coquel, A priori and a posteriori analysis of finite volume discretizations of Darcy's equations. Numer. Math. **96**(1) (2003), 17–42.

[2] A. Alonso, Error estimators for a mixed method. Numer. Math. **74** (1996), 385–395.

[3] I. Babuška and G. N. Gatica, On the mixed finite element method with Lagrange multipliers. Numer. Meth. Partial Diff. Eq. **19**(2) (2003), 192–210.

[4] I. Babuška and W. C. Rheinboldt, Error estimates for adaptive finite element computations. SIAM J. Numer. Anal. **15** (1978), 736–754.

[5] T. Belytschko and T. Blacker, Enhanced derivative recovery through least square residual penalty. Appl. Num. Math. **14** (1994), 55–68.

[6] D. Braess and R. Verfürth, A posteriori error estimators for the Raviart-Thomas element. SIAM J. Numer. Anal. **33**(6) (1996), 2431–2444.

[7] J. H. Brandts, Superconvergence and a posteriori error estimation for triangular mixed finite elements. Numer. Math. **68**(3) (1994), 311–324.

[8] J. H. Brandts and Y. Chen, An alternative to the least-squares mixed finite element method for elliptic problems. In Numerical Mathematics and Advanced Applications, eds. M. Feistauer, V. Dolejši, P. Knobloch, K. Najzar, Springer, Berlin, pp. 169–175, 2004.

[9] F. Brezzi and M. Fortin, Mixed and Hybrid Finite Element Methods. Springer-Verlag, New York, 1991.

[10] Z. Cai, R. Lazarov, T. A. Manteuffel and S. F. McCormick, First-order system least squares for second-order partial differential equations: Part I. SIAM J. Num. Anal. **31**(6) (1994), 1785–1799.

[11] Z. Cai, T. A. Manteuffel, S. F. McCormick and S. V. Parter, First-order system least squares (FOSLS) for planar linear elasticity: Pure traction problem. SIAM J. Num. Anal. **35** (1998), 320–335.

[12] C. Carstensen, A posteriori error estimate for the mixed finite element method. Math. Comp. **66**(218) (1997), 465–476.

[13] C. Carstensen and S. Bartels, Each averaging technique yields reliable a posteriori error control in FEM on unstructured grids. I Low order conforming, non-conforming and mixed FEM. Math. Comp. **71**(239) (2002), 945–969.

[14] C. Carstensen and G. Dolzmann, A posteriori error estimates for mixed FEM in elasticity. Numer. Math. **81** (1998), 187–209.

[15] I. Ekeland and R. Temam, Convex Analysis and Variational Problems. North-Holland, New York, 1976.

[16] G. Gatica and M. Maischak, A posteriori error estimates for the mixed finite element method with Lagrange multipliers. Numer. Meth. Partial Diff. Eq., to appear.

[17] V. Girault and P.-A. Raviart, Finite Element Methods for Navier-Stokes Equations. Theory and Algorithms. Springer-Verlag, New York, 1986.

[18] M. Křížek and P. Neittaanmäki, Superconvergence phenomenon in the finite element method arising from averaging of gradients. Numer. Math. **45** (1984), 105–116.

[19] P. Ladeveze and D. Leguillon, Error estimate procedure in the finite element method and applications. SIAM J. Numer. Anal. **20** (1983), 485–509.

[20] M. Lonsing and R. Verfürth, A posteriori error estimators for mixed finite element methods in linear elasticity. Numer. Math. **97** (2004), 757–778.

[21] A. I. Pehlivanov, G. F. Carey and R. D. Lazarov, Least-squares mixed finite elements for second order elliptic problems. SIAM J. Num. Anal. **31**(5) (1994), 1368–1377.

[22] W. Prager and J. L. Synge, Approximations in elasticity based on the concept of functions space. Quart. Appl. Math. **5** (1947), 241–269.

[23] P.-A. Raviart and J.-M. Thomas, A mixed finite element for second order elliptic problems. In Mathematical Aspects of Finite Element Methods, eds. I. Galligani and E. Magenes, Springer, Berlin, pp. 292–315, 1977.

[24] S. Repin, A posteriori error estimation for variational problems with uniformly convex functionals. Math. Comp. **69**(230) (2000), 481–500.

[25] S. Repin, Two-sided estimates of deviation from exact solutions of uniformly elliptic equations. Amer. Math. Soc. Transl. Series 2 **209** (2003), 143–171.

[26] S. Repin, S. Sauter and A. Smolianski, A posteriori error estimation for the Dirichlet problem with account of the error in approximation of boundary conditions. Computing **70**(3) (2003), 205–233.

[27] S. Repin, S. Sauter and A. Smolianski, A posteriori error estimation for the Poisson equation with mixed Dirichlet/Neumann boundary conditions. J. Comput. Appl. Math. **164/165** (2004), 601–612.

[28] S. Repin, S. Sauter and A. Smolianski, A posteriori estimation of dimension reduction errors for elliptic problems on thin domains. SIAM J. Num. Anal. **42**(4) (2004), 1435–1451.

[29] J. E. Roberts and J.-M. Thomas, Mixed and Hybrid Methods. In Handbook of Numerical Analysis, II, eds. P. G. Ciarlet and J.-L. Lions, North-Holland, Amsterdam, pp. 523–639, 1991.

[30] N.-E. Wiberg, F. Abdulwahab and S. Ziukas, Enhanced superconvergent patch recovery incorporating equilibrium and boundary conditions. Int. J. Num. Meth. Engng. **37** (1994), 3417–3440.

[31] B. I. Wohlmuth and R. H. W. Hoppe, A comparison of a posteriori error estimators for mixed finite element discretizations by Raviart-Thomas elements. Math. Comp. **68**(228) (1999), 1347–1378.

[32] Z. Zhang and A. Naga, A new finite element gradient recovery method: superconvergence property. SIAM J. Sci. Comput. **26** (2005), 1192–1213.

[33] O. C. Zienkiewicz and J. Z. Zhu, A simple error estimator and adaptive procedure for practical engineering analysis. Int. J. Num. Meth. Engng. **24** (1987), 337–357.

[34] O. C. Zienkiewicz and J. Z. Zhu, The superconvergent patch recovery and a posteriori error estimates. I. The recovery technique. Int. J. Num. Meth. Engng. **33** (1992), 1331–1364.